UNITÉ DE RECHERCHE
INRIA-RENNES

Institut National
de Recherche
en Informatique
et en Automatique

Domaine de Voluceau
Rocquencourt
B.P. 105
78153 Le Chesnay Cedex
France
Tél.:(1) 39 63 55 11

Rapports de Recherche

1992

25ème anniversaire

N° 1736

*Programme 1*
*Architectures parallèles, Bases de données, Réseaux et Systèmes distribués*

# AGRÉGATION FAIBLE DES PROCESSUS DE MARKOV ABSORBANTS

James LEDOUX
Gerardo RUBINO
Bruno SERICOLA

Septembre 1992

*RR-1736*

# IRISA

INSTITUT DE RECHERCHE EN INFORMATIQUE
ET SYSTEMES ALEATOIRES

Campus Universitaire de Beaulieu
35042 - RENNES CEDEX FRANCE
Tél. : 99 84 71 00 - Télex : UNIRISA 950 473 F
Télécopie : 99 38 38 32

# Agrégation faible des processus de Markov absorbants

James Ledoux, Gerardo Rubino, Bruno Sericola.

Programme I

Publication Interne No 666, Juillet 1992, 30 pages.

## Résumé

Nous caractérisons les conditions dans lesquelles un processus markovien absorbant (à temps discret ou continu) d'espace d'état fini, peut être transformé en un nouveau processus, dit agrégé, qui soit encore markovien mais dont les états sont les classes d'une partition donnée de l'espace d'état originel. Pour cela, la notion pertinente est la distribution quasi-stationnaire associée aux processus absorbants étudiés ici. Elle permet d'établir un lien explicite avec le cas irréductible traité dans [5]. Nous sommes alors en mesure de calculer l'ensemble de toutes les distributions initiales du processus originel conduisant à un processus agrégé markovien au moyen d'un algorithme fini. Enfin, il est montré que l'étude d'un processus à temps continu se ramène à celle d'une chaîne à temps discret en utilisant la technique d'uniformisation.

**Mots clés:** Processus de Markov, Agrégation faible, Distribution Quasi-Stationnaire, Uniformisation.

# Weak Lumpability of Absorbing Markov Processes

## Abstract

We characterize the conditions under which an absorbing markovian finite process (in discrete or continuous time) can be transformed into a new aggregated process, conserving the markovian property, whose states are elements of a given partition of the original state space. To obtain this characterization, the suitable notion is the quasi-stationary distribution associated with absorbing processes. It allows the absorbing case to be related to the irreducible one. We are able to calculate the set of all initial distributions of the starting model leading to an aggregated homogeneous Markov process by means of a finite algorithm. Finally, it is shown that the continuous time case can always be reduced to the discrete one using the uniformization technique.

**Keywords:** Markov Processes, Weak Lumpability, Quasi-Stationary Distribution, Uniformization.

# Table des matières

**1 Introduction** **5**

**2 Notations et définitions** **5**

**3 Distribution quasi-stationnaire** **7**
3.1 Cas discret . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 8
3.2 Cas continu . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 10

**4 Agrégation forte et faible dans le cas discret** **11**
4.1 Résultats préliminaires . . . . . . . . . . . . . . . . . . . . . . . . 11
4.2 Caractérisation ponctuelle . . . . . . . . . . . . . . . . . . . . . . . 13
4.3 Etude des distributions initiales . . . . . . . . . . . . . . . . . . . 14
4.4 Lien avec le cas irréductible . . . . . . . . . . . . . . . . . . . . . 18
4.5 Exemple . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 20

**5 Agrégation forte et faible dans le cas continu** **21**
5.1 Caractérisation de l'agrégation faible. . . . . . . . . . . . . . . . . 22
5.2 Lien avec le cas discret . . . . . . . . . . . . . . . . . . . . . . . . 26

# 1 Introduction

Depuis fort longtemps, de nombreux chercheurs se sont intéressés aux transformations fonctionnelles de processus de Markov. Une des raisons essentielles est qu'elles correspondent souvent à la seule "réponse" observable ou d'intérêt d'un système caractérisé (ou modélisé) par une évolution markovienne. La fonctionnelle étudiée ici, consiste pour l'utilisateur à ne se préoccuper (ou ne disposer) que de l'information fournie par des agrégats d'états du modèle initial de son système. Ainsi le processus d'intérêt est le processus dit agrégé, dont l'espace d'état est constitué de classes d'états du modèle originel. Il est alors très intéressant, pour conserver toute la puissance des outils d'analyse des processus de Markov, de pouvoir affirmer si le processus transformé est encore markovien. Des conditions nécessaires et suffisantes ont été obtenues par Kemeny et Snell [3], Rubino et Sericola [5],[6],[7] pour des processus irréductibles. Ce rapport propose de telles caractérisations pour des processus absorbants, modèles que l'on rencontre par exemple pour des systèmes à durée de vie finie. Il se décompose de la façon suivante. La Section 2 introduit toutes les notations utilisées dans les sections suivantes, de même que la définition formelle du processus agrégé. La Section 3 présente en détail la notion de distribution quasi-stationnaire associée aux processus étudiés ici. Les deux résultats énoncés donnent une interprétation asymptotique bien connue de cette distribution. Cependant elle a été précisée pour les chaînes périodiques et les preuves sont à notre connaissance originales. La Section 4 est dévolue aux processus markoviens évoluant en temps discret. Elle apporte une condition nécessaire et suffisante pour obtenir un processus agrégé markovien. Nous montrons également que la distribution quasi-stationnaire fait toujours partie des vecteurs de probabilités initiales du processus originel conduisant à un agrégé markovien. Enfin on met en évidence un lien entre le cas absorbant et irréductible permettant d'utiliser l'algorithme fini de [6]. Cette construction est illustrée sur un exemple. Finalement, le résultat principal de la Section 5 établit que le cas des processus évoluant en temps continu se réduit à celui des chaînes à temps discret grâce à la technique d'uniformisation.

# 2 Notations et définitions

Cette Section introduit les principales notations et des lemmes préliminaires. On considère un processus markovien homogène absorbant $X$, à temps discret ou continu. On supposera que ce processus ne possède qu'un unique état absorbant, ceci sans aucune perte de généralité comme on le verra dans la Section 4.

- L'espace d'état est fini et noté $E = \{1, 2, ..., N\}$.
- On note $\mathcal{B} = \{B(0), B(1), B(2), ..., B(M)\}$ une partition quelconque de l'espace d'état $E$ où $B(0)$ ne contient que l'état absorbant du processus noté 1. Le cardinal de la classe $B(l)$ est noté $n(l)$. Pour plus de clarté, on suppose les états de $E$ ordonnés

de la façon suivante:

$$\begin{array}{rcl} B(0) & = & \{1\}, \\ B(1) & = & \{n(0)+1, \ldots, n(0)+n(1)\}, \\ & \vdots & \\ B(l) & = & \{n(0)+ \cdots + n(l-1)+1, \ldots, n(0)+ \cdots + n(l)\}, \\ & \vdots & \\ B(M) & = & \{n(0)+ \cdots + n(M-1)+1, \ldots, N\}. \end{array}$$

- On suppose que l'ensemble $T$ des états transitoires du processus forme une classe irréductible et ouverte. Cette dernière propriété signifie simplement qu'il existe au moins un état $j \in T$ communiquant avec l'état absorbant 1.

- Au processus $X$, on associe le processus stochastique agrégé $Y$ sur l'espace d'état $F = \{0, 1, 2, ..., M\}$, défini par :

$$Y_t = l \iff X_t \in B(l) \text{ pour tout } t.$$

- Par convention, tous les vecteurs utilisés sont des vecteurs lignes. Les vecteurs colonnes seront représentés par l'opérateur $(.)^*$. Si $v$ est un vecteur, la notation $v > 0$ indique que toutes ses composantes sont strictement positives. Le vecteur $e_i$ symbolise le $i^{ème}$ vecteur de la base canonique de $\mathbb{R}^N$. Un vecteur dont toutes les composantes sont égales à 1 sera noté simplement 1, sa dimension étant définie par le contexte. Une convention similaire est adoptée pour le vecteur dont toutes les composantes sont nulles. La matrice diagonale de terme diagonal générique $v(i)$ sera notée $diag(v)$. Sa dimension sera définie par le contexte, de même que pour la matrice identité qui est notée $I$.

- On notera $\mathcal{A}$ l'ensemble des vecteurs de probabilité de dimension $N$ et $\mathcal{A}^T$ le sous-ensemble de $\mathcal{A}$ constitué des distributions de probabilité de support $T$, i.e. $\mathcal{A}^T = \{\alpha \in \mathcal{A} / \sum_{i \in T} \alpha(i) = 1\}$.

- On rencontrera constamment des égalités de la forme $\mathbb{P}_u(f(X)) = \mathbb{P}_v(g(X))$, où $X$ est une famille de chaînes de Markov de probabilités de transition fixées. La partie gauche (resp. droite) de l'égalité représente alors la probabilité de l'événement $f(X)$ (resp. $g(X)$) pour l'élément de la famille de chaînes de Markov spécifié par la distribution initiale $u$ (resp. $v$) (avec $u, v \in \mathcal{A}$).

- Pour tout $l \in F$, $\alpha_{B(l)}$ note la restriction de $\alpha$ à la classe $B(l)$. Ce vecteur ligne a pour dimension $n(l)$ et sa $k^{ième}$ composante est la $(n(0) + \cdots + n(l-1) + k)^{ième}$ composante de $\alpha$. On a ainsi définit l'opérateur linéaire

$$R_l : \quad \mathcal{A} \longrightarrow [\,0, 1\,]^{n(l)}$$
$$\alpha \mapsto \alpha_{B(l)}.$$

Cette définition sera également utilisée pour la restriction de $\alpha \in \mathcal{A}$ à tout sous-ensemble d'états de $E$.

Réciproquement, à tout vecteur de probabilité $\beta$ de dimension $n(l)$, on peut associer le vecteur $R_l^{-1}.\beta = \gamma \in \mathcal{A}$, avec $\gamma(i) = 0$ si $i \notin B(l)$ et $\gamma(i) = \beta(i-j)$ si $i \in B(l)$ et $j = n(0) + \cdots + n(l-1)$. Là encore, cette construction pourra être utilisée pour plonger tout vecteur de probabilité de dimension inférieure à $N$ dans $\mathcal{A}$.

- Pour $\alpha \in \mathcal{A}$ et $B \in \mathcal{B}$, on notera $\alpha^B$ le vecteur de $\mathcal{A}$ défini par

$$\alpha^B(i) = \begin{cases} \dfrac{\alpha(i)}{\sum_{j \in B} \alpha(j)} & \text{si } i \in B, \\ 0 & \text{sinon.} \end{cases}$$

Son existence est conditionnée par la relation $\alpha_B 1^* \neq 0$.

Les propriétés suivantes sont énoncées sans preuve.

**Lemme 2.1** *[6]*
1. *Tout élément $\alpha \in \mathcal{A}$ s'écrit de manière unique comme combinaison convexe des $\alpha^{B(l)}$ lorsqu'ils sont définis. C'est à dire:*

$$\alpha = \sum_{\substack{l \in F, \\ \alpha_{B(l)} \neq 0}} \alpha_{B(l)} 1^* \alpha^{B(l)}$$

2. *Soit $l \in F$, $x_1, \ldots, x_n$ $n$ vecteurs de $\mathcal{A}$, $\lambda_1, \ldots, \lambda_n$ $n$ réels positifs tels que $\sum_{i=1}^n \lambda_i = 1$ et $K$ une constante positive quelconque, alors*

$$\left(K \sum_{i=1}^n \lambda_i x_i\right)^{B(l)} = \left(\sum_{i=1}^n \lambda_i x_i\right)^{B(l)} = \sum_{\substack{i/1 \leq i \leq n, \\ (x_i)_{B(l)} \neq 0}} \frac{\lambda_i (x_i)_{B(l)} 1^*}{\sum_{j=1}^n \lambda_j (x_j)_{B(l)} 1^*} (x_i)^{B(l)}$$

*lorsque* $\left(\sum_{i=1}^n \lambda_i x_i\right)^{B(l)}$ *est défini, c'est à dire lorsque* $\left(\sum_{i=1}^n \lambda_i x_i\right)_{B(l)} \neq 0.$

# 3 Distribution quasi-stationnaire

La notion de distribution quasi-stationnaire pour les processus markoviens absorbants a été développée par Darroch et Seneta [1],[2],[9]. Ce vecteur de probabilité est l'analogue pour une matrice sous-stochastique irréductible, du vecteur de probabilité stationnaire associé à une matrice stochastique irréductible. Cette distribution jouera le même rôle que le vecteur de probabilité stationnaire dans l'agrégation des chaînes de Markov homogènes irréductibles. Nous allons d'abord traiter le cas du temps discret puis le cas du temps continu.

## 3.1 Cas discret

La matrice de transition $P$ de la chaîne peut être décomposée de la façon suivante

$$P = \left( \begin{array}{c|c} 1 & 0 \\ \hline q^* & Q \end{array} \right).$$

De même, le vecteur de la distribution initiale peut s'exprimer sous la forme $\alpha = (\alpha_{\{1\}}, \alpha_T)$.

L'hypothèse d'irréductibilité de la classe $T$ entraîne l'irréductibilité de la matrice $Q$. Nous pouvons alors appliquer le théorème de Perron-Frobenius [9, Th.1.5,page 22] à la matrice sous-stochastique $Q$. Il assure l'existence d'une unique valeur propre réelle $\rho$, strictement positive et majorant en module toutes les autres valeurs propres de la matrice $Q$. Elle est appellée rayon spectral de la matrice $Q$. Ce théorème établit également l'existence d'un unique vecteur de probabilité $v > 0$ (resp. $w^* > 0$), vecteur propre à gauche (resp. vecteur propre à droite) associé à $\rho$ de la matrice $Q$.

Pour tout $n \in \mathbb{N}^*$, nous définissons alors le vecteur de probabilité $\pi_n$ sur $T$ suivant

- si $Q$ est apériodique:

$$\pi_n(j) \overset{\text{déf}}{=} \mathbb{P}_\alpha\{X_n = j | X_n \in T\} = \frac{\alpha_T Q^n e_j{}^*}{\alpha_T Q^n 1^*}. \tag{1}$$

- si $Q$ est périodique:

$$\pi_n(j) \overset{\text{déf}}{=} \frac{\sum_{k=1}^{n} \alpha_T r^k Q^k e_j{}^*}{\sum_{k=1}^{n} \alpha_T r^k Q^k 1^*} \qquad \text{où } r = \frac{1}{\rho}. \tag{2}$$

Nous allons étudier le comportement des distributions (1),(2) quand $n$ devient grand. Pour cela nous nous ramenons à l'étude de ces limites lorsque $Q$ est remplacée par une certaine matrice stochastique $\overline{Q}$. En effet, si nous posons

$$\overline{Q} = \frac{1}{\rho} W^{-1} Q W \tag{3}$$

avec $W = diag(w)$, cette matrice $\overline{Q}$ est positive, stochastique et possède la même topologie de 0 que $Q$. Elle est donc irréductible, apériodique (resp. périodique) si et seulement si la matrice $Q$ est irréductible, apériodique (resp. périodique). Nous sommes en mesure d'établir le résultat énoncé dans [1] pour une matrice $Q$ apériodique et démontré en utilisant les propriétés spectrales de cette matrice. Ici, grâce à la matrice $\overline{Q}$ définie en (3), nous allons nous appuyer sur les théorèmes classiques décrivant le comportement asymptotique des chaînes de Markov.

**Théorème 3.1** *Si la matrice positive $Q$ est irréductible alors pour toute distribution initiale $\alpha \in \mathcal{A}$ et pour tout $j \in T$ nous avons*

$$\lim_{n \longrightarrow \infty} \pi_n(j) = v(j) > 0,$$

*où $v$ est l'unique vecteur de probabilité solution de $vQ = \rho v$ avec $\rho$ rayon spectral de $Q$. Cette limite est indépendante de la distribution initiale $\alpha$ et est appellée la distribution quasi-stationnaire de la chaîne de Markov absorbante $X$.*

**Démonstration.** La matrice $\overline{Q}$ définie en (3) est irréductible par hypothèse sur $Q$. Elle admet donc un unique vecteur de probabilité invariant $\pi$. Alors pour tout $\alpha \in \mathcal{A}$ et $j \in T$ nous avons dans le cas où $Q$ est apériodique:

$$\begin{aligned}
\lim_{n \longrightarrow \infty} \pi_n(j) &= \lim_{n \longrightarrow \infty} \frac{\alpha_T W \overline{Q}^n W^{-1} e_j^*}{\alpha_T W \overline{Q}^n W^{-1} 1^*} \\
&= \frac{\alpha_T W 1^* \pi W^{-1} e_j^*}{\alpha_T W 1^* \pi W^{-1} 1^*} \\
&= \frac{\pi W^{-1} e_j^*}{\pi W^{-1} 1^*} > 0.
\end{aligned}$$

Pour le cas périodique, un calcul analogue avec l'expression (2), conduit au même vecteur limite. Posons

$$v = \frac{\pi W^{-1}}{\pi W^{-1} 1^*},$$

il est alors facile de vérifier que $vQ = \rho v$ où $\rho$ est le rayon spectral de la matrice $Q$. Donc $v$ est l'unique vecteur propre à gauche de $Q$ associé à $\rho$, strictement positif et tel que $v 1^* = 1$. Enfin ce vecteur est bien indépendant de la distribution initiale $\alpha$. □

Le vecteur de probabilité $v$ représente bien un régime stationnaire pour les distributions (1) ou (2), dans le sens où si la distribution initiale de la chaîne vaut $(0, v)$, nous avons alors pour tout $n \in \mathbb{N}^*$ et $Q$ apériodique (le cas périodique est analogue),

$$\begin{aligned}
\pi_{n+1}(j) &= \frac{v Q^{n+1} e_j^*}{v Q^{n+1} 1^*} = \frac{\rho v Q^n e_j^*}{\rho v Q^n 1^*} \quad (\text{car } vQ = \rho v) \\
&= \pi_n(j) \\
&= \frac{\rho^n v e_j^*}{\rho^n v 1^*} = v(j).
\end{aligned}$$

C'est à dire:

$$\forall n \in \mathbb{N}^*, \;\; \pi_{n+1}(j) = \pi_n(j) = v(j).$$

En résumé, pour une chaîne de Markov homogène absorbante avec un ensemble des états transitoires formant une classe irréductible, $\pi_n$ converge vers une distribution de probabilité $v$ indépendamment de la distribution initiale $\alpha$. Ce vecteur $v > 0$ est l'unique vecteur de probabilité solution de $vQ = \rho v$.

## 3.2 Cas continu

Comme pour le cas discret, nous décomposons la matrice des taux de transition $A$ du processus de Markov en

$$A = \left(\begin{array}{c|c} 0 & 0 \\ \hline g^* & G \end{array}\right).$$

Les hypothèses sur la classe $T$ se traduisent par le fait que $g \neq 0$ et que $G$ est une matrice irréductible. Le semi-groupe des probabilités de transitions est donnée par $P_t = e^{At}$. Notons $Q_t$ la matrice $e^{Gt}$ obtenue en suprimant la première ligne et la première colonne de $P_t$. Nous définissons alors le vecteur de probabilité

$$\pi_t(j) \stackrel{\text{déf}}{=} \mathbb{P}_\alpha(X_t = j / X_t \in T) = \frac{\alpha_T Q_t e_j^*}{\alpha_T Q_t 1^*}. \tag{4}$$

La matrice irréductible $G$ est telle que $G(i,j) \geq 0$ pour $i \neq j$. Nous pouvons alors utiliser le résultat de [9, Th. 2.6, page 46]. Il établit l'existence pour cette matrice, d'une valeur propre $\rho$ réelle, simple, auquelle est associée un unique vecteur propre $v > 0$ à gauche (resp $w^* > 0$ à droite). Cette valeur propre, là encore nommée rayon spectral de la matrice $G$, est telle que $\rho > Re(\lambda)$, où $\lambda$ est une valeur propre (éventuellement complexe) distincte de $\rho$ et de partie réelle notée $Re(\lambda)$.
Posons

$$\overline{G} = W^{-1}(G - \rho I)W \quad \text{où } W = diag(w). \tag{5}$$

Il est simple de vérifier que la matrice $\overline{G}$ possède la même topologie de 0 que $G$, que $\overline{G}(i,j) \geq 0$ pour $i \neq j$ et qu'elle satisfait $\overline{G}1^* = 0$. Elle est ainsi irréductible si et seulement si $G$ est irréductible. Nous pouvons démontrer l'analogue du Théorème 3.1 pour les processus à temps continu, en utilisant les propriétés asymptotiques de la famille de processus markoviens irréductibles de générateur $\overline{G}$.

**Théorème 3.2** *[2] Si la matrice $G$ est irréductible, alors pour toute loi initiale $\alpha \in \mathcal{A}$ et pour tout $j \in T$*

$$\lim_{t \longrightarrow +\infty} \pi_t(j) = v(j) > 0,$$

*où $v$ est l'unique vecteur de probabilité solution de $vG = \rho v$ avec $\rho$ le rayon spectral de la matrice $G$. Cette limite est indépendante de $\alpha$et est appellée la distribution quasi-stationnaire du processus de Markov absorbant $X$.*

**Démonstration.** La matrice $\overline{G}$ définie en (5) permet d'écrire $Q_t$ sous la forme

$$Q_t = e^{Gt} = e^{\rho t}\, W\, e^{\overline{G}t}\, W^{-1}.$$

Nous savons que $\overline{G}$ admet un unique vecteur de probabilité stationnaire $\pi$ tel que

$$\lim_{t \longrightarrow +\infty} e^{\overline{G}t} \;=\; 1^*\pi, \quad \text{avec} \quad \pi\overline{G} = 0 \text{ et } \pi 1^* = 1.$$

Alors pour tout $\alpha \in \mathcal{A}$ et pour tout $j \in T$ nous avons

$$\begin{aligned}
\lim_{t \longrightarrow +\infty} \pi_t(j) &= \lim_{t \longrightarrow +\infty} \frac{\alpha_T W e^{\overline{G}t} W^{-1} e_j{}^*}{\alpha_T W e^{\overline{G}t} W^{-1} 1^*} \\
&= \frac{\alpha_T W 1^* \pi W^{-1} e_j{}^*}{\alpha_T W 1^* \pi W^{-1} 1^*} \\
&= \frac{\pi W^{-1} e_j{}^*}{\pi W^{-1} 1^*} > 0.
\end{aligned}$$

Là encore, on vérifie aisément que le vecteur de probabilité

$$z = \frac{\pi W^{-1}}{\pi W^{-1} 1^*} > 0,$$

est l'unique vecteur propre à gauche associé au rayon spectral $\rho$ de la matrice $G$, strictement positif tel que $z1^* = 1$. Enfin ce vecteur est bien indépendant de la distribution initiale $\alpha$. □

On peut faire la même remarque qu'en fin de section précédente, concernant le rôle de la distribution quasi-stationnaire pour le vecteur $\pi_t$ défini en (4).

# 4 Agrégation forte et faible dans le cas discret

## 4.1 Résultats préliminaires

Soit $X = (X_n)_{n \geq 0}$ une chaîne de Markov homogène sur l'espace d'état $E$. On suppose que la partition $\mathcal{B}$ est fixée. $X$ est donnée par sa matrice des probabilités de transition $P$ et par sa distribution de probabilité initiale $\alpha$. On notera $(\alpha, P)$ cette chaîne de Markov lorsque ce sera nécessaire. Les éléments de $P$ seront notés $P(i,j)$. On notera $agr(\alpha, P, \mathcal{B})$ la chaîne agrégée construite par rapport à la partition $\mathcal{B}$ à partir de $(\alpha, P)$.

On considèrera souvent la famille de toutes les chaînes de Markov sur l'espace d'état $E$ possédant la même matrice des probabilités de transition $P$. Cette famille sera notée $(\,.\,, P)$. En général, la distribution de probabilité initiale figurera en indice de la mesure de probabilité. Par exemple, $\mathbb{P}_\alpha(X_n \in B)$ représente la probabilité qu'après $n$ transitions, la chaîne de Markov $(\alpha, P)$ se trouve dans l'un des états du sous-ensemble $B$ de $E$.

On notera $P(i,B)$ la probabilité de passer en une transition de l'état $i$ à l'un des états du sous-ensemble $B$ de $E$, c'est à dire, $P(i,B) = \sum_{j\in B} P(i,j)$. Si l'on considère la décomposition par blocs de la matrice $P$ relativement à la partition $\mathcal{B}$, alors on notera $P_{B(i)B(j)}$ le bloc $n(i) \times n(j)$ correspondant aux transitions de $B(i)$ vers $B(j)$.

**Définition 4.1** *Une suite $(C_0, C_1, \ldots, C_j)$ de sous-ensembles de $E$ sera dite possible pour la distribution initiale $\alpha$ si $\mathbb{P}_\alpha(X_0 \in C_0, X_1 \in C_1, \ldots, X_j \in C_j) > 0$. En particulier, si $B \in \mathcal{B}$, $(B)$ est une suite possible pour $\alpha$ ssi $\alpha_B \neq 0$.*

Ainsi la suite $(C_0, C_1, \ldots, C_j)$ est possible pour $\alpha$, si la probabilité pour que la chaîne $(\alpha, P)$ se trouve aux instants $0, 1, 2, \ldots, j$ dans les sous-ensembles $(C_0, C_1, \ldots, C_j)$ n'est pas nulle. Une remarque immédiate est que si cette suite contient la classe absorbante $B(0)$ alors elle devient stationnaire de terme général égal à $B(0)$. Donc toute suite possible se terminant par un élément de $\mathcal{B}$ distinct de $B(0)$ ne contient pas $B(0)$.

**Définition 4.2** *Soit $\alpha \in \mathcal{A}$ et $(C_0, C_1, \ldots, C_j)$ une suite d'éléments de $\mathcal{B}$ possible pour $\alpha$, on définit le vecteur $f(\alpha, C_0, C_1, \ldots, C_j) \in \mathcal{A}$ récursivement par:*

$$\begin{aligned} f(\alpha, C) &= \alpha^C \\ f(\alpha, C_0, C_1, \ldots, C_k) &= (f(\alpha, C_0, C_1, \ldots, C_{k-1})P)^{C_k}. \end{aligned}$$

Par exemple, $f(\alpha, B, C, D) = ((\alpha^B P)^C P)^D$.
Nous aurons parfois à utiliser simultanément des fonctions $f$ construites à partir de matrices des probabilités de transition distinctes. Dans ce cas, la matrice invoquée sera placée en indice de la fonction $f$ correspondante.

Nous rappelons les propriétés suivantes valables pour une chaîne de Markov homogène quelconque.

**Lemme 4.3** *[7] Pour tout $k \geq 0$ on a*

$$\mathbb{P}_\alpha(X_{n+1} \in B / X_n \in C_n, \ \ldots, \ X_{n-k} \in C_{n-k}) = \mathbb{P}_\beta(X_1 \in B) \ \text{ pour tout } n \geq k$$

$$\text{où } \beta = f(\alpha P^{n-k}, C_{n-k}, \ \ldots, C_n).$$

*Pour $k = n$, le vecteur $\beta = f(\alpha, C_0, C_1, \ldots, C_n)$ est la distribution de $X_n$ sachant $\{X_0 \in C_0, \ \ldots, \ X_n \in C_n\}$. C'est à dire:*

$$\beta(i) = \mathbb{P}_\alpha(X_n = i / X_0 \in C_0, \ \ldots, \ X_n \in C_n)$$

Enfin le dernier résultat de cette section est un lemme technique.

**Lemme 4.4** *[8] La fonction $v \mapsto \mathbb{P}_v(X_j \in B / X_{j-1} \in C_{j-1}, \ \ldots, X_{j-k} \in C_{j-k})$ pour $j$, $k$, $C_{j-1}, \ldots, C_{j-k}$, $B$ fixés, est continue sur $\{v \in \mathcal{A}$ tel que $(C_{j-k}, \ldots, C_{j-1})$ est possible pour $v\}$.*

## 4.2 Caractérisation ponctuelle

Nous analysons maintenant le processus $Y = agr(\alpha, P, \mathcal{B})$. Pour tout $B \in \mathcal{B}$, on notera $\mathcal{A}(\alpha, B)$ le sous-ensemble de $\mathcal{A}$ constitué par les distributions de probabilité de la forme $f(\alpha, C_1, \ldots, C_j, B)$, c'est à dire:

$$\mathcal{A}(\alpha,B) \stackrel{\text{déf}}{=} \{\ \beta \in \mathcal{A}\ /\ \exists j \geq 0 \text{ et une suite } (C_1,\ldots,C_j,B) \text{ possible pour } \alpha, \\ \text{réduite à } (B) \text{ si } j=0, \text{ tels que } \beta = f(\alpha, C_1,\ldots,C_j,B)\ \}.$$

Il est facile de vérifier que pour tout $\alpha \in \mathcal{A}$ telle que $\alpha_T 1^* \neq 0$ et pour tout $B \in \mathcal{B}$, l'ensemble $\mathcal{A}(\alpha, B)$ n'est pas vide; c'est une conséquence de l'irréductibilité de $Y$ sur $F \setminus \{0\}$ et du fait que $T$ est ouverte. La même démarche que [6, Th 2.2] conduit à une condition nécessaire et suffisante pour que $Y$ soit une chaîne de Markov homogène, après avoir remarqué que toute suite possible poura, de dernier élément $B(l)$ avec $l \neq 0$, ne contient jamais $B(0)$ et que pour tout $\beta \in \mathcal{A}(\alpha, B(0))$

$$\mathbb{P}_\beta(X_1 \in B(m)) = 0 \text{ pour } m \neq 0 \quad \text{et} \quad \mathbb{P}_\beta(X_1 \in B(0)) = 1.$$

**Théorème 4.5** *La chaîne $Y = agr(\alpha, P, \mathcal{B})$ est une chaîne de Markov homogène si et seulement si $\forall l \in F \setminus \{0\}, \forall m \in F$, la probabilité $\mathbb{P}_\beta(X_1 \in B(m))$ a la même valeur pour tout $\beta \in \mathcal{A}(\alpha, B(l))$. Cette valeur commune est la probabilité de transition de l'état $l$ vers l'état $m$ pour la chaîne $Y$.*

**Remarque.** Pour l'étude de l'agrégation des chaînes absorbantes, il est équivalent de considérer les chaînes avec une classe $B(0)$ ne contenant qu'un seul état absorbant. En effet, si $X = (\alpha, P)$ possède $1, 2, \ldots, n(0)$ (avec $n(0) > 1$ ) comme états absorbants, on peut construire un nouveau processus $X' = (\alpha', P')$ ne contenant qu'un seul état absorbant noté 1, possédant la même classe des états transitoires (renumérotés de 2 à $N - n(0)$ dans $X'$) et tel que pour $i, j = 2, \ldots, N - n(0) + 1$,

$$\begin{aligned}
P'(i,j) &= P(n(0)-1+i, n(0)-1+j), \\
P'(j,1) &= P(n(0)-1+j, B(0)), \\
P'(1,j) &= \begin{cases} 0 \text{ ssi } j \neq 1 \\ 1 \text{ autrement}, \end{cases} \\
\alpha'(i) &= \alpha(n(0)-1+i), \\
\alpha'(1) &= \alpha_{B(0)} 1^*.
\end{aligned}$$

La partition $\mathcal{B}'$ pour $X'$ se déduit directement de la partition induite par $\mathcal{B}$ sur le sous-ensemble $T$. On peut alors montrer que $X$ et $X'$ sont équivalents au sens de l'agrégation, i.e. $agr(\alpha, P, \mathcal{B})$ est markovien dès que l'agrégé de $X'$ l'est. Pour cela, il suffit de voir que $\forall l \neq 0, \forall \beta = f_P(\alpha, C_0, \ldots, C_k, B(l)), \forall m \in F$,

$$\mathbb{P}_\beta(X_1 \in B(m)) = \mathbb{P}_\delta(X_1' \in B'(m)) \qquad \text{où } \delta = f_{P'}(\alpha', C_0', \ldots, C_k', B'(l)).$$

L'existence de chaînes de Markov satisfaisant le Théorème 4.5 se déduit de leur existence dans le cas irréductible [3] avec le corollaire suivant.

**Corollaire 4.6** *Soit une chaîne de Markov homogène $(\alpha, P)$ irréductible dont la chaîne agrégée selon une partition $\mathcal{B}$ est markovienne. Supposons qu'il existe un élément $B$ de cette partition tel que $E \setminus B$ forme une classe irréductible. Alors, la chaîne de Markov absorbante déduite de $(\alpha, P)$ en rendant les états de $B$ absorbants, i.e. en modifiant la matrice $P$ par: $\forall i \in B$, $P(i,j) = 0$ $\forall j \in E \setminus B$ et $P(i,i) = 1$, possède une chaîne agrégée selon $\mathcal{B}$, markovienne.*

## 4.3 Etude des distributions initiales

Nous allons maintenant étudier de façon générale l'ensemble des distributions initiales $\alpha$ conduisant à une chaîne de Markov homogène pour $Y = agr(\alpha, P, \mathcal{B})$. Nous noterons cet ensemble $\mathcal{A}_{\mathcal{M}}$, c'est à dire

$$\mathcal{A}_{\mathcal{M}} \stackrel{\text{déf}}{=} \{\alpha \in \mathcal{A} /\ Y = agr(\alpha, P, \mathcal{B}) \text{ est une chaîne de Markov homogène}\}.$$

Là encore, on aura à utiliser simultanément de tels ensembles générés par différentes familles de chaînes de Markov. Dans ce cas, nous utiliserons la notation $\mathcal{A}_{\mathcal{M}}(P)$ pour l'ensemble associé à la famille $(., P)$.

Nous pouvons adapter sans difficulté le Lemme 3.4 de [8] donnant les propriétés suivantes de l'ensemble $\mathcal{A}_{\mathcal{M}}$.

**Lemme 4.7** *Si $\alpha \in \mathcal{A}_{\mathcal{M}}$ alors $\alpha P^j \in \mathcal{A}_{\mathcal{M}}$ pour tout $j \in \mathbb{N}$; si $k$ est tel que $\alpha_{B(k)} 1^* \neq 0$ alors $\alpha^{B(k)} \in \mathcal{A}_{\mathcal{M}}$.*

Il n'est pas très difficile de constater que pour étudier l'ensemble $\mathcal{A}_{\mathcal{M}}$, il suffit de considérer son sous-ensemble, noté $\mathcal{A}_{\mathcal{M}}^T$, formé par les distributions sur $E$ de support l'ensemble $T$, ceci grâce à l'égalité suivante

$$\mathcal{A}_{\mathcal{M}} = \lambda_1 \{e_1\} + \lambda_T \mathcal{A}_{\mathcal{M}}^T \quad \text{où} \quad \lambda_T, \lambda_1 \geq 0 \text{ et } \lambda_T + \lambda_1 = 1. \tag{6}$$

Nous nous restreignons donc à l'étude de l'ensemble

$$\mathcal{A}_{\mathcal{M}}^T = \{\alpha \in \mathcal{A}^T / Y = agr(\alpha, P, \mathcal{B}) \text{ est une chaîne de Markov homogène}\}.$$

**Lemme 4.8** *Soit $r = 1/\rho$ avec $\rho$ le rayon spectral de $Q$. Si $\alpha \in \mathcal{A}_{\mathcal{M}}^T$ alors pour tout $n \geq 1$:*

$$\left(0, \frac{\alpha_T Q^n}{\alpha_T Q^n 1^*}\right) \in \mathcal{A}_{\mathcal{M}}^T \quad \text{et} \quad \left(0, \frac{\sum_{k=1}^{n} \alpha_T Q^k r^k}{\sum_{k=1}^{n} \alpha_T Q^k r^k 1^*}\right) \in \mathcal{A}_{\mathcal{M}}^T.$$

**Démonstration.** Par le Lemme 4.7 nous savons que $\alpha P^n \in \mathcal{A}_{\mathcal{M}}$ pour tout $n \in \mathbb{N}$. Mais l'égalité (6) nous dit que ceci est équivalent à $(\alpha P^n)^T \in \mathcal{A}_{\mathcal{M}}^T$. Comme

$$P^n = \left(\begin{array}{c|c} 1 & 0 \\ \hline q_n^* & Q^n \end{array}\right)$$

on a

$$\alpha P^n = (\alpha_{\{1\}} + \alpha_T {q_n}^*, \alpha_T Q^n). \tag{7}$$

Et par définition nous obtenons finalement

$$(\alpha P^n)^T = \left(0, \frac{(\alpha P^n)_T}{(\alpha P^n)_T 1^*}\right) = \left(0, \frac{\alpha_T Q^n}{\alpha_T Q^n 1^*}\right).$$

Maintenant, notons $\gamma_n$ le vecteur suivant

$$\gamma_n = \left(0, \frac{\sum_{k=1}^{n} \alpha_T Q^k r^k}{\sum_{k=1}^{n} \alpha_T Q^k r^k 1^*}\right).$$

Par le Lemme 4.3, nous avons pour tout $l \in F \setminus \{0\}$ et $m \in F$,

$$\mathbb{P}_{\gamma_n}(X_{n+1} \in B(m)/X_n \in B(l),\ X_{n-1} \in C_{n-1},\ \ldots, X_0 \in C_0) = \mathbb{P}_{\gamma_n'}(X_1 \in B(m))$$

$$\text{où } \gamma_n' = f(\gamma_n,\ C_0,\ \ldots,\ C_{n-1},\ B(l)).$$

Si on note $\alpha_k'$ le vecteur $f\left(\left(0, \frac{\alpha_T Q^k}{\alpha_T Q^k 1^*}\right),\ C_0,\ \ldots,\ C_{n-1},\ B(l)\right)$ alors $\alpha_k' = f(\alpha P^k,\ C_0, \ldots,\ C_{n-1},\ B(l))$, et on montre que

$$\gamma_n' = \sum_{k=1}^{n} \frac{K_k}{K} \alpha_k'$$

$$\text{où } K_k = \frac{1}{\sum_{k=1}^{n} \alpha_T Q^k r^k 1^*} \left(\alpha_T Q^k r^k\right)_{C_0} P_{C_0 C_1} \ldots P_{C_{n-1} B(l)} 1^* \quad \text{et } K = \sum_{k=1}^{n} K_k.$$

On obtient alors

$$\mathbb{P}_{\gamma_n'}(X_1 \in B(m)) = \sum_{k=1}^{n} \frac{K_k}{K} \mathbb{P}_{\alpha_k'}(X_1 \in B(m)). \tag{8}$$

Or:

$$\begin{aligned}
\mathbb{P}_{\alpha_k'}(X_1 \in B(m)) &= \mathbb{P}_{\alpha P^k}(X_{n+1} \in B(m)/X_n \in B(l),\ \ldots,\ X_0 \in C_0) \\
&= \mathbb{P}_{\alpha}(X_{n+k+1} \in B(m)/X_{n+k} \in B(l),\ \ldots,\ X_k \in C_0) \\
&= \mathbb{P}_{\alpha}(X_{j+1} \in B(m)/X_j \in B(l)) \\
& \quad \forall j \text{ tel que } \mathbb{P}_{\alpha}(X_j \in B(l)) > 0, \text{ puisque } \alpha \in \mathcal{A}_{\mathcal{M}}^T.
\end{aligned}$$

Finalement, pour tout $j$ tel que $\mathbb{P}_{\alpha}(X_j \in B(l)) > 0$ on a avec (8)

$$\mathbb{P}_{\gamma_n'}(X_1 \in B(m)) = \mathbb{P}_{\alpha}(X_{j+1} \in B(m)/X_j \in B(l))$$

qui est donc constante sur $\mathcal{A}(\gamma_n, B(l))$. Ainsi, pour tout $n \geq 1$, $\gamma_n \in \mathcal{A}_{\mathcal{M}}^T$. □

**Théorème 4.9** *Si $\mathcal{A}_{\mathcal{M}}^T \neq \emptyset$ et si $\widehat{P}$ désigne la matrice de transition de la chaîne de Markov homogène $Y = agr(\alpha, P, \mathcal{B})$ alors cette matrice $\widehat{P}$ est la même pour tout $\alpha \in \mathcal{A}_{\mathcal{M}}^T$. De plus, la distribution de probabilité quasi-stationnaire $v$ associée à la classe des états transitoires $T$ est telle que $(0, v) \in \mathcal{A}_{\mathcal{M}}^T$.*

**Démonstration.** Soit $\alpha \in \mathcal{A}^T$ tel que $Y = agr(\alpha, P, \mathcal{B})$ est une chaîne de Markov homogène (i.e. $\alpha \in \mathcal{A}_{\mathcal{M}}^T$). Désignons par $\widehat{P}(l,m)$ la probabilité de transition de $l$ vers $m$ pour $Y$. On a alors pour tout $l \neq 0$ et $\forall k$ tel que $\mathbb{P}_\alpha(X_k \in B(l)) > 0$,

$$\begin{aligned}
\widehat{P}(l,m) &= \mathbb{P}_\alpha(X_{k+1} \in B(m)/X_k \in B(l)) \\
&= \mathbb{P}_{(\alpha P^k)^{B(l)}}(X_1 \in B(m)) \\
&= \mathbb{P}_{\left(0, \frac{\alpha_T Q^k}{\alpha_T Q^k 1^*}\right)^{B(l)}}(X_1 \in B(m)) \quad \text{par (7)} \\
&= \mathbb{P}_{\left(0, \frac{\alpha_T Q^k}{\alpha_T Q^k 1^*}\right)}(X_1 \in B(m)/X_0 \in B(l)).
\end{aligned}$$

Supposons d'abord que $Q$ est apériodique. On a par le Lemme 4.8:

$$\alpha \in \mathcal{A}_{\mathcal{M}}^T \Longrightarrow \left(0, \frac{\alpha_T Q^k}{\alpha_T Q^k 1^*}\right) \in \mathcal{A}_{\mathcal{M}}^T \quad \text{pour tout } k \geq 1.$$

Soit $k_0$ suffisament grand pour que $\forall k \geq k_0$, $(\alpha_T Q^k)_{B(l)} 1^* > 0$ (un tel $k_0$ existe car $Q$ est irréductible et apériodique). En faisant tendre $k$ vers $+\infty$ et grâce au Lemme 4.4, on obtient:

$$\widehat{P}(l,m) = \mathbb{P}_{(0,v)}(X_1 \in B(m)/X_0 \in B(l))$$

qui ne dépend pas de $\alpha$.

Dans le cas périodique nous avons, toujours par le lemme 4.8:

$$\forall n \geq 1 \quad \left(0, \frac{\sum_{k=1}^{n} \alpha_T Q^k r^k}{\sum_{k=1}^{n} \alpha_T Q^k r^k 1^*}\right) \in \mathcal{A}_{\mathcal{M}}^T.$$

Notons $\beta_n$ ce dernier vecteur. Soit $n_0$ suffisament grand pour que $\forall n \geq n_0$, $\sum_{k=1}^{n} (\alpha_T r^k Q^k)_{B(l)}$ $1^* > 0$ (un tel $n_0$ existe car $X$ est irréductible). Pour simplifier l'écriture, posons pour $k = 1, \ldots, n$

$$\gamma_k = \frac{(\alpha_T Q^k r^k)_{B(l)} 1^*}{\sum_{k=1}^{n} (\alpha_T Q^k r^k)_{B(l)} 1^*}.$$

L'élément $\widehat{P}(l,m)$ peut alors s'écrire de la façon suivante:

$$\begin{aligned}
\widehat{P}(l,m) &= \sum_{k=1}^{n} \gamma_k \widehat{P}(l,m) \qquad \text{car } \sum_{k=1}^{n} \gamma_k = 1 \\
&= \sum_{\substack{k:1\le k\le n,\\ \mathbb{P}_\alpha(X_k\in B(l))>0}} \gamma_k \mathbb{P}_{\left(0,\frac{\alpha_T Q^k}{\alpha_T Q^k 1^*}\right)}(X_1 \in B(m)/X_0 \in B(l)) \\
&= \sum_{\substack{k:1\le k\le n,\\ \mathbb{P}_\alpha(X_k\in B(l))>0}} \gamma_k \mathbb{P}_{\left(0,\frac{\alpha_T Q^k}{\alpha_T Q^k 1^*}\right)^{B(l)}}(X_1 \in B(m)) \\
&= \mathbb{P}_\Gamma(X_1 \in B(m))
\end{aligned}$$

$$\text{où } \Gamma = \sum_{\substack{k:1\le k\le n,\\ \mathbb{P}_\alpha(X_k\in B(l))>0}} \gamma_k \left(0, \frac{\alpha_T Q^k}{\alpha_T Q^k 1^*}\right)^{B(l)}.$$

En utilisant le Lemme 2.1, $\Gamma$ peut s'écrire:

$$\Gamma = \left(0, \frac{\sum_{k=1}^{n} \alpha_T Q^k r^k}{\sum_{k=1}^{n} \alpha_T Q^k r^k 1^*}\right)^{B(l)} = \beta_n{}^{B(l)}.$$

On obtient finalement:

$$\widehat{P}(l,m) = \mathbb{P}_{\beta_n}(X_1 \in B(m)/X_0 \in B(l)).$$

En faisant tendre $n$ vers $+\infty$ et toujours grâce au Lemme 4.4, on obtient là encore

$$\widehat{P}(l,m) = \mathbb{P}_{(0,v)}(X_1 \in B(m)/X_0 \in B(l))$$

qui ne dépend pas de $\alpha$. On tire de cette dernière expression que $\forall l \neq 0$ , $\forall m \in F$,

$$\widehat{P}(l,m) = \sum_{i\in B(l)} v^{B(l)}(i)\ P(i, B(m)). \tag{9}$$

$\square$

Ainsi, si $\mathcal{A}_{\mathcal{M}}^T \neq \emptyset$, la chaîne $Y = agr(\alpha, P, \mathcal{B})$, pour $\alpha \in \mathcal{A}_{\mathcal{M}}^T$, a pour matrice de transition la matrice notée $\widehat{P}$, dont l'élément $(l,m)$ est décrit par (9) et pour distribution initiale $\varphi_B(\alpha)$, si $\varphi_B$ est la fonction

$$\alpha \mapsto (\ \alpha_{B(0)}.1^*, \alpha_{B(1)}.1^*, \ldots, \alpha_{B(M)}.1^*\ ).$$

**Définition 4.10** *On dira que la famille de chaînes de Markov $(., P)$ est faiblement agrégeable selon la partition $\mathcal{B}$ si et seulement si l'ensemble $\mathcal{A}^T_{\mathcal{M}} \neq \emptyset$.*
*Pour tout $\alpha \in \mathcal{A}_{\mathcal{M}}$, la chaîne agrégée $Y = agr(\alpha, P, \mathcal{B})$ est une chaîne de Markov homogène de matrice des probabilités de transition $\widehat{P}$ donnée dans le Théorème 4.9.*

Un cas particulier important est celui où tout les $\alpha \in \mathcal{A}$ donnent une chaîne $Y = agr(\alpha, P, \mathcal{B})$ homogène et markovienne. On dit alors que la famille $(., P)$ est fortement agrégeable ou agrégeable par rapport à la partition $\mathcal{B}$. En utilisant le même type d'argument que dans [3], on peut alors caractériser ces familles par

**Théorème 4.11** *$(., P)$ est fortement agrégeable selon la partition $\mathcal{B}$ si et seulement si pour toute paire de sous-ensembles $D, B \in \mathcal{B}$, $P(d, B)$ a la même valeur pour tout $d \in D$. Cette valeur commune est la probabilité de transition de $D$ vers $B$ pour la chaîne $Y$.*

**Remarque.** Grâce à la construction de la matrice $P'$ page 13 et l'égalité (6), nous obtenons dans le cas où le processus absorbant de matrice $P$ admet $n(0) > 1$ états absorbants,

$$\mathcal{A}_{\mathcal{M}}(P) = \sum_{l=1}^{n(0)} \lambda_l e_l + \lambda_{n(0)+1} R_T^{-1}.\mathcal{A}^T_{\mathcal{M}}(P'), \tag{10}$$
$$\text{avec } \lambda_l \geq 0 \text{ et } \sum_{l=1}^{n(0)+1} \lambda_l = 1.$$

## 4.4 Lien avec le cas irréductible

La matrice $P$ des probabilités de transition peut s'écrire sous la forme

$$P = \begin{pmatrix} 1 & 0 \\ (I-Q).1^* & Q \end{pmatrix}.$$

Nous pouvons alors construire la matrice $P^{(v)}$ suivante:

$$P^{(v)} = \begin{pmatrix} 0 & v \\ (I-Q).1^* & Q \end{pmatrix}$$

où $v$ est le vecteur de probabilité quasi-stationnaire associé à $P$. Cette matrice $P^{(v)}$ est la matrice des probabilités de transition d'une chaine de Markov irréductible.

**Théorème 4.12** *Si $\mathcal{A}_{\mathcal{M}}(P^{(v)})$ note l'ensemble des distributions initiales $\alpha$ telles que $Y = agr(\alpha, P^{(v)}, \mathcal{B})$ est markovien, nous avons*

$$\alpha \in \mathcal{A}_{\mathcal{M}}(P) \Longleftrightarrow \alpha \in \mathcal{A}_{\mathcal{M}}(P^{(v)}).$$

*De plus, pour tout $l = 1, \ldots, M$ et $m = 0, \ldots, M$ on a $\widehat{P}(l, m) = \widehat{P^{(v)}}(l, m)$.*

**Démonstration.** Nous rappelons que $f_P$ (resp. $f_{P^{(v)}}$) représente la fonction $f$ de la Définition 4.2 lorsque $P$ (resp. $P^{(v)}$) est invoquée pour sa construction. De même $X$ (resp. $X'$) représente une chaîne de Markov de matrice des probabilités de transition $P$ (resp. $P^{(v)}$).

Si $\alpha \in \mathcal{A}_{\mathcal{M}}(P^{(v)})$, nous avons par définition que pour tout $l = 0, \ldots, M$ et pour tout $\beta' = f_{P^{(v)}}(\alpha, C_0, \ldots, B(l))$

$$\mathbb{P}_{\beta'}(X_1' \in B(m)) = \sum_{i \in B(l)} \beta'(i) P^{(v)}(i, B(m)) = \widehat{P^{(v)}}(l, m), \quad m = 0, \ldots, M. \tag{11}$$

Montrons alors que $\alpha \in \mathcal{A}_{\mathcal{M}}(P)$. Par construction de la matrice $P^{(v)}$, nous avons les deux assertions suivantes pour tout $l \neq 0$:

$$\begin{aligned} f_{P^{(v)}}(\alpha, C_0, \ldots, B(l)) &= f_P(\alpha, C_0, \ldots, B(l)), \\ \forall i \in B(l): \ P^{(v)}(i, B(m)) &= P(i, B(m)) \text{ pour tout } m = 0, \ldots, M; \end{aligned}$$

d'où pour tout $m = 0, \ldots, M$

$$\begin{aligned} \mathbb{P}_\beta(X_1 \in B(m)) &= \sum_{i \in B(l)} \beta(i) P(i, B(m)) \\ &= \sum_{i \in B(l)} \beta'(i) P^{(v)}(i, B(m)) \\ &= \widehat{P^{(v)}}(l, m). \end{aligned}$$

La caractérisation du Théorème 4.5 est ainsi satisfaite.

Inversement, supposons que $\alpha \in \mathcal{A}_{\mathcal{M}}(P)$. Si $l = 0$: tout vecteur $\beta' = f_{P^{(v)}}(\alpha, C_0, \ldots, B(l))$ se réduit à $e_1$. Auquel cas nous avons

$$\mathbb{P}_{\beta'}(X_1' \in B(m)) = \begin{cases} (e_1 P^{(v)})_{B(m)}.1^* = v_{B(m)}.1^* & \text{pour } m \neq 0, \\ 0 & m = 0. \end{cases}$$

Ainsi cette probabilité est constante et ne dépend que de $m$. Elle est égale par définition à $\widehat{P^{(v)}}(0, m)$.

Si $l \neq 0$: supposons d'abord que le vecteur $\beta'$ fasse intervenir au moins une fois la classe $B(0)$. Nous pouvons écrire ce vecteur sous la forme suivante:

$$\beta' = f_{P^{(v)}}(\alpha, C_0, \ldots, C_{j-1}, B(0), C_{j+1}, \ldots, C_n, B(l))$$

où $j$ est le plus grand entier compris entre 0 et $n$ tel que la suite $C_{j+1}, \ldots, C_n, B(l)$ ne contienne pas $B(0)$ (avec la convention que si $j = n$ alors la suite est réduite à $B(l)$). En utilisant la définition récursive 4.2 de la fonction $f$, $\beta'$ s'écrit encore

$$\beta' = f_{P^{(v)}}\left(f_{P^{(v)}}(\alpha, C_0, \ldots, C_{j-1}, B(0)) P^{(v)}, C_{j+1}, \ldots, C_n, B(l)\right).$$

Comme $B(0)$ est réduite à un seul état, on a $f_{P^{(v)}}(\alpha, C_0, \ldots, C_{j-1}, B(0)) P^{(v)} = (0, v)$. Nous obtenons alors

$$\beta' = f_{P^{(v)}}((0, v), C_{j+1}, \ldots, C_n, B(l)).$$

Mais dans la construction de ce dernier vecteur, seuls les coefficients de la matrice $Q$ interviennent. Ce qui permet de réécrire $\beta'$ sous la forme

$$\beta' = f_P((0,v), C_{j+1}, \ldots, C_n, B(l)).$$

On en déduit alors que pour tout $\beta'$ et $m = 0, \ldots, M$

$$\begin{aligned} \mathbb{P}_{\beta'}(X_1' \in B(m)) &= \mathbb{P}_\gamma(X_1 \in B(m)) \quad \text{avec } \gamma = f_P((0,v), C_{j+1}, \ldots, C_n, B(l)) \\ &= \hat{P}(l,m) \quad \text{car } (0,v) \in \mathcal{A}_{\mathcal{M}}(P). \end{aligned}$$

Enfin supposons que $\beta'$ ne contient pas $B(0)$, alors le vecteur $\beta' = f_P(\alpha, C_0, \ldots, B(l))$ car seule la matrice $Q$ intervient. On en déduit

$$\mathbb{P}_{\beta'}(X_1' \in B(m)) = \hat{P}(l,m).$$

Ce qui termine la démonstration. □

**Remarque.** Il est simple de voir que le vecteur des probabilités stationnaires $\pi'$ de la matrice $P^{(v)}$ peut s'écrire $\pi' = \lambda(0,v) + (1-\lambda)e_1$ avec $\lambda \geq 0$. Ce qui nous donne directement avec (9) l'égalité des probabilités de transition des chaînes agregées anoncée dans le théorème précédent.

Ce dernier théorème nous permet d'utiliser directement l'algorithme fini obtenu dans [6] pour les chaînes de Markov irréductibles.

## 4.5 Exemple

Supposons que la chaîne de Markov possède la matrice des probabilités de transition suivante:

$$P = \left(\begin{array}{cc|ccc|c} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ \hline 1/12 & 1/12 & 1/6 & 1/6 & 1/3 & 1/6 \\ 1/24 & 1/24 & 3/8 & 1/4 & 1/6 & 1/8 \\ 1/24 & 1/24 & 1/8 & 1/4 & 1/6 & 3/8 \\ \hline 1/14 & 1/14 & 3/14 & 3/14 & 3/14 & 3/14 \end{array}\right)$$

et soit $\mathcal{B} = \{B(0), B(1), B(2)\}$, $B(0) = \{1,2\}$, $B(1) = \{3,4,5\}$ et $B(2) = \{6\}$. Alors avec la construction de la page 13, nous pouvons nous en tenir à étudier la chaîne possédant un seul état absorbant de matrice des probabilités de transition:

$$P' = \left(\begin{array}{c|ccc|c} 1 & 0 & 0 & 0 & 0 \\ \hline 1/6 & 1/6 & 1/6 & 1/3 & 1/6 \\ 1/12 & 3/8 & 1/4 & 1/6 & 1/8 \\ 1/12 & 1/8 & 1/4 & 1/6 & 3/8 \\ \hline 1/7 & 3/14 & 3/14 & 3/14 & 3/14 \end{array}\right).$$

La sous-matrice $Q$ est donnée par

$$Q = \begin{pmatrix} 1/6 & 1/6 & 1/3 & 1/6 \\ 3/8 & 1/4 & 1/6 & 1/8 \\ 1/8 & 1/4 & 1/6 & 3/8 \\ 3/14 & 3/14 & 3/14 & 3/14 \end{pmatrix}.$$

Nous calculons la distribution quasi-stationnaire associée à la sous-matrice $Q$: son rayon spectral vaut 37/42 et le vecteur propre à gauche associé $v = (1/4, 1/4, 1/4, 1/4)$. Ce qui nous donne $v^{B(1)} = (1/3, 1/3, 1/3, 0)$ et $v^{B(2)} = (0, 0, 0, 1)$. Nous pouvons alors construire la matrice $P'^{(v)}$ du Théorème 4.12:

$$P'^{(v)} = \left(\begin{array}{c|ccc|c} 0 & 1/4 & 1/4 & 1/4 & 1/4 \\ \hline 1/6 & 1/6 & 1/6 & 1/3 & 1/6 \\ 1/12 & 3/8 & 1/4 & 1/6 & 1/8 \\ 1/12 & 1/8 & 1/4 & 1/6 & 3/8 \\ \hline 1/7 & 3/14 & 3/14 & 3/14 & 3/14 \end{array}\right).$$

En appliquant l'algorithme donné dans [6], on obtient les sommets du polyèdre $\mathcal{A}_{\mathcal{M}}(P'^{(v)}) = \mathcal{A}_{\mathcal{M}}(P')$

$$(1,0,0,0,0), (0,1/3,1/3,1/3,0), (0,0,0,0,1).$$

Pour la matrice $P$ initiale, grâce à la relation (10), $\mathcal{A}_{\mathcal{M}}(P)$ est le polyèdre convexe défini par les quatre sommets

$$(1,0,0,0,0,0), (0,1,0,0,0,0), (0,0,1/3,1/3,1/3,0), (0,0,0,0,0,1).$$

Suivant le Théorème 4.12 et (9), la matrice de transition $\widehat{P}$ associée aux différentes chaînes agrégées markovienne issues de $\mathcal{A}_{\mathcal{M}}(P)$ est

$$\widehat{P} = \begin{pmatrix} 1 & 0 & 0 \\ 1/9 & 2/3 & 2/9 \\ 1/7 & 9/14 & 3/14 \end{pmatrix}.$$

# 5 Agrégation forte et faible dans le cas continu

Soit $X = (X_t)_{t\geq 0}$ un processus markovien homogène absorbant d'espace d'état $E$. On désigne par $A$ sa matrice des taux de transition. De même que dans le cas discret, $(\alpha, A)$ désigne le processus de Markov absorbant de distribution initiale $\alpha$. La quantité $A(i, B(m))$ désigne le taux de transition de l'état i vers la classe d'état $B(m)$ i.e $A(i, B(m)) = \sum_{j\in B(m)} A(i,j)$. La quantité $e^{At}(i, B)$ représente la probabilité que le processus soit dans le sous-ensemble $B$ de $E$ à l'instant $t$ sachant qu'il démarre dans l'état $i$, i.e $e^{At}(i, B) = \sum_{j\in B} e^{At}(i,j)$.

## 5.1 Caractérisation de l'agrégation faible.

Nous allons suivre la même démarche que dans le cas discret. Tout d'abord nous avons l'analogue en temps continu des définitions et lemmes de la Section 4.1.

**Définition 5.1** *Une suite $(C_0, C_1, \ldots, C_j)$ de sous-ensembles de $E$ sera dite possible pour la distribution initiale $\alpha$ si pour tout $0 = t_0 < t_1 < \cdots < t_j$, $\mathbb{P}_\alpha(X_{t_0} \in C_0, X_{t_1} \in C_1, \ldots, X_{t_j} \in C_j) > 0$. En particulier, si $B \in \mathcal{B}$, $(B)$ est une suite possible pour $\alpha$ si $\alpha_B \neq 0$.*

Comme dans le cas discret, une remarque immédiate est que toute suite possible se terminant par un élément de $\mathcal{B}$ distinct de $B(0)$, ne contient jamais $B(0)$.

Soit $(C_0, \ldots, C_j)$ une suite d'éléments de $\mathcal{B}$. Soit $\mathcal{A}(C_0, \ldots, C_j) \overset{\text{déf}}{=} \{\gamma \in \mathcal{A} /$ la suite $(C_0, \ldots, C_j)$ est possible pour $\gamma\}$. De la même manière que dans [5], on peut montrer que $\mathcal{A}(C_0, \ldots, C_j)$ est convexe.

**Définition 5.2** *Soit $\alpha \in \mathcal{A}$ et $(C_0, C_1, \ldots, C_j)$ une suite d'éléments de $\mathcal{B}$ possible pour $\alpha$, on définit, pour tout $0 = t_0 < t_1 < \cdots < t_j$, le vecteur de probabilité $h(\alpha, (C_0, t_0), (C_1, t_1), \ldots, (C_j, t_j))$ récursivement par:*

$$\begin{aligned} h(\alpha, (C, 0)) &= \alpha^C \\ h(\alpha, (C_0, t_0), (C_1, t_1)) &= \left(h(\alpha, (C_0, t_0))e^{At_1}\right)^{C_1} \\ &\textit{et pour } k \geq 2: \\ h(\alpha, (C_0, t_0), (C_1, t_1), \ldots, (C_k, t_k)) &= \left(h(\alpha, (C_0, t_0), (C_1, t_1), \ldots, (C_{k-1}, t_{k-1}))e^{A(t_k - t_{k-1})}\right)^{C_k}. \end{aligned}$$

Par exemple, $h(\alpha, (B, 0), (C, t_1), (D, t_2)) = \left((\alpha^B e^{At_1})^C e^{A(t_2 - t_1)}\right)^D$.

Le lemme suivant est à la base de tout ce qui suit.

**Lemme 5.3** *[5] Pour tout $k \geq 0$, on a:*

$$\mathbb{P}_\alpha(X_{t_n + t} \in B / X_{t_{n-k}} \in C_{n-k}, \ \ldots, \ X_{t_n} \in C_n) = \mathbb{P}_\beta(X_t \in B) \ \ \textit{pour tout } n \geq k$$

$$\textit{où } \beta = h(\alpha e^{At_{n-k}}, (C_{n-k}, 0), \ \ldots, (C_n, t_n)).$$

*Le vecteur $\beta$ représente la distribution de $X_{t_n}$ sachant que $X_{t_{n-k}} \in C_{n-k}, \ \ldots, \ X_{t_n} \in C_n$.*

Nous aurons également besoin du résultat de continuité suivant:

**Lemme 5.4** *[6] La fonction $v \mapsto \mathbb{P}_v(X_{t_j} \in B / X_{t_{j-1}} \in C_{j-1}, \ldots, X_{t_{j-k}} \in C_{j-k})$ pour $j$, $k$, $C_{j-1}, \ldots, C_{j-k}$, $B$, $0 < t_{j-k} < \cdots < t_j$ fixés, est continue sur $\{v \in \mathcal{A}$ tel que $(C_{j-k}, \ldots, C_{j-1})$ est possible pour $v\}$.*

Comme nous l'avons fait dans le cas discret, nous allons étudier le processus $Y = agr(\alpha, A, \mathcal{B})$. On notera par $\mathcal{C}(\alpha, B)$ l'ensemble suivant:

$$\mathcal{C}(\alpha, B) \stackrel{\text{déf}}{=} \{\ \beta \in \mathcal{A}\ /\ \exists j \geq 0 \text{ et une suite } (C_1, \ldots, C_j, B) \text{ possible pour } \alpha, \text{ réduite à } (B) \text{ si } j = 0, \text{ tels que } \beta = h(\alpha, (C_1, 0), (C_2, t_1), \ldots, (C_j, t_{j-1}), (B, t_j)), \text{ pour tout } 0 = t_0 < t_1 < \cdots < t_j.\}$$

Cet ensemble est l'équivalent de l'ensemble $\mathcal{A}(\alpha, B)$ dans le cas discret. Le théorème suivant est une adaptation directe du [5, Th. 4.11] en utilisant les mêmes remarques que dans Théorème 4.5.

**Théorème 5.5** *Le processus $Y = agr(\alpha, A, \mathcal{B})$ est un processus markovien homogène si et seulement si $\forall l \in F \setminus \{0\}, m \in F$ et $\forall t \geq 0$, la probabilité $\mathbb{P}_\beta(X_t \in B(m))$ a la même valeur pour tout $\beta \in \mathcal{C}(\alpha, B(l))$. Cette valeur commune est la probabilité que le processus $Y$ soit dans l'état $m$ à l'instant $t$ sachant qu'il est dans l'état $l$ à l'instant 0.*

En fait, en suivant [5, Th. 4.12], on peut montrer le résultat suivant.

**Théorème 5.6** *Le processus $Y = agr(\alpha, A, \mathcal{B})$ est un processus markovien homogène si et seulement si $\forall l \in F \setminus \{0\}, m \in F$, la quantité $\sum_{i \in B(l)} \beta(i) A(i, B(m))$ est la même pour tout $\beta \in \mathcal{C}(\alpha, B(l))$. Cette valeur commune est le taux de transition de l'état $l$ vers l'état $m$ pour $Y$ lorsque $l \neq m$.*

On notera $\mathcal{C}_\mathcal{M}$ l'ensemble des distributions initiales $\alpha$ conduisant à un processus markovien homogène, c'est à dire:

$$\mathcal{C}_\mathcal{M} \stackrel{\text{déf}}{=} \{\alpha \in \mathcal{A} / Y = agr(\alpha, A, \mathcal{B}) \quad \text{est un processus markovien homogène}\}.$$

Soit $\mathcal{C}_\mathcal{M}^T$ le sous-ensemble de $\mathcal{C}_\mathcal{M}$ des distributions de support l'ensemble $T$, nous avons une égalité du même type que (6)

$$\mathcal{C}_\mathcal{M} = \lambda_1 \{e_1\} + \lambda_T \mathcal{C}_\mathcal{M}^T \qquad \text{où } \lambda_T, \lambda_1 \geq 0 \text{ et } \lambda_T + \lambda_1 = 1. \tag{12}$$

En conservant la même démarche que pour le cas discret et en décomposant la matrice $A$ comme dans la Section 3.2, on obtient le résultat suivant.

**Lemme 5.7** *Si $\alpha \in \mathcal{C}_\mathcal{M}$ alors pour tout $s \geq 0$:*

$$\alpha e^{As} \in \mathcal{C}_\mathcal{M} \quad et \quad \left(0, \frac{\alpha_T e^{Gs}}{\alpha_T e^{Gs} \mathbb{1}^*}\right) \in \mathcal{C}_\mathcal{M}^T$$

**Démonstration.** Soit $\alpha \in \mathcal{C}_\mathcal{M}$. En utilisant le Lemme 5.3, il est simple de vérifier que:

$$\mathbb{P}_\alpha(X_{t_{n+1}+s} \in B / X_{t_n+s} \in D,\ X_{t_{n-1}+s} \in C_{n-1},\ \ldots,\ X_s \in C_0)$$
$$= \mathbb{P}_{\alpha e^{As}}(X_{t_{n+1}} \in B / X_{t_n} \in D,\ X_{t_{n-1}} \in C_{n-1},\ \ldots,\ X_0 \in C_0).$$

Par hypothèse, la partie gauche de cette égalité est indépendante de $n$, de $s$ et de la suite $(C_0, C_1, \ldots, C_{n-1})$ considérée (les classes $B$ et $D$ étant fixées). Donc, $\alpha e^{As} \in \mathcal{C}_\mathcal{M}$ pour tout $s \geq 0$.

Grace à ce qui précède, si $\alpha \in \mathcal{C}_{\mathcal{M}}$ nous avons que pour tout $s \geq 0$ que $\alpha e^{As} \in \mathcal{C}_{\mathcal{M}}$. L'égalité (12) nous dit que ceci est équivalent à $(\alpha e^{As})^T \in \mathcal{C}^T_{\mathcal{M}}$. Comme

$$e^{As} = \left( \begin{array}{c|c} 1 & 0 \\ \hline g_s^* & e^{Gs} \end{array} \right)$$

on a

$$\alpha e^{As} = (\alpha_{\{1\}} + \alpha_T g_s^*, \alpha_T e^{Gs}). \tag{13}$$

Par définition on a

$$\left(\alpha e^{As}\right)^T = \left(0, \frac{(\alpha e^{As})_T}{(\alpha e^{As})_T 1^\bullet}\right) = \left(0, \frac{\alpha_T e^{Gs}}{\alpha_T e^{Gs} 1^\bullet}\right).$$

□

Nous pouvons maintenant démontrer l'analogue du Théorème 4.9 pour les processus à temps continu.

**Théorème 5.8** *Si $\mathcal{C}^T_{\mathcal{M}} \neq \emptyset$ et si $\widehat{A}$ désigne la matrice de transition du processus de Markov homogène $Y = agr(\alpha, P, \mathcal{B})$ alors la matrice $\widehat{A}$ est la même pour tout $\alpha \in \mathcal{C}^T_{\mathcal{M}}$. De plus la distribution de probabilité quasi-stationnaire $v$ associée à la matrice $G$ est telle que $(0, v) \in \mathcal{C}^T_{\mathcal{M}}$.*

**Démonstration.** Soit $\alpha \in \mathcal{A}$ tel que $Y = agr(\alpha, A, \mathcal{B})$ est un processus markovien homogène de matrice des taux de transition $\widehat{A}$. On a alors pour $l \neq 0$ et $t \geq 0$,

$$\begin{aligned} e^{\widehat{A}s}(l,m) &= \mathbb{P}_\alpha(X_{t+s} \in B(m) / X_t \in B(l)) \\ &= \mathbb{P}_{(\alpha e^{At})^{B(l)}}(X_s \in B(m)) \\ &= \mathbb{P}_{\left(0, \frac{\alpha_T e^{Gt}}{\alpha_T e^{Gt} 1^\bullet}\right)^{B(l)}}(X_s \in B(m)) \\ &= \mathbb{P}_{\left(0, \frac{\alpha_T e^{Gt}}{\alpha_T e^{Gt} 1^\bullet}\right)}(X_s \in B(m) / X_0 \in B(l)). \end{aligned}$$

En faisant tendre $t$ vers $+\infty$, on obtient grâce au Lemme 5.4

$$e^{\widehat{A}s}(l,m) = \mathbb{P}_{(0,v)}(X_s \in B(m) / X_0 \in B(l)).$$

Ce qui donne

$$\begin{aligned} \widehat{A}(l,m) &= \lim_{s \to 0} \frac{1}{s} \mathbb{P}_{(0,v)}(X_s \in B(m) / X_0 \in B(l)) \text{ pour } l \neq m \\ \widehat{A}(l,l) &= \lim_{s \to 0} \frac{1}{s} \left( \mathbb{P}_{(0,v)}(X_s \in B(m) / X_0 \in B(l)) - 1 \right). \end{aligned} \tag{14}$$

On tire de (14) que $\forall l \neq 0$ , $\forall m \neq l$

$$\widehat{A}(l,m) = \sum_{i \in B(l)} v^{B(l)}(i) \; A(i, B(m)). \tag{15}$$

$\square$

Ainsi, si $\mathcal{C}_{\mathcal{M}}^T \neq \emptyset$, le processus $Y = agr(\alpha, A, \mathcal{B})$ a pour matrice de taux de transition la matrice dont l'élément $(l, m)$ est décrit par (15) et pour loi initiale $\varphi_{\mathcal{B}}(\alpha)$.

On conserve les mêmes définitions pour les notions *"fortement agrégeable"* (ou *"agrégeable"*) et *"faiblement agrégeable"*. La notion d'agrégation forte étant un cas particulier de celle d'agrégation faible, on ne parlera dans cette section que du cas général.

**Remarque.** Par une construction du même type que celle de la page 13, le cas des processus à multiples états absorbants se réduit à l'étude d'un processus avec un unique état absorbant, équivalent au processus original au sens de l'agrégation.

Nous allons maintenant nous intéresser au calcul de l'ensemble $\mathcal{C}_{\mathcal{M}}^T$. Pour cela, on introduit les notations suivantes.

- Pour $i \in F \setminus \{0\}$, on notera $\widetilde{A}_i$ la matrice $(n(i), M+1)$ définie par: $\widetilde{A}_i(j,k) \stackrel{\text{déf}}{=} A(n(0) + \cdots + n(i-1) + j, B(k))$, $1 \leq j \leq n(i)$, $k \in F$
- On notera $\widehat{A}_0, \ldots, \widehat{A}_j, \ldots, \widehat{A}_M$ les $(M+1)$ lignes successives de la matrice $\widehat{A}$ des taux de transition définie par:
$$\widehat{A}_j \stackrel{\text{déf}}{=} (R_j.v^{B(j)})\widetilde{A}_j \quad \text{pour } j \in F \setminus \{0\} \quad \text{et} \quad \widehat{A}_0 \stackrel{\text{déf}}{=} (0, \ldots, 0). \tag{16}$$
Par la relation (15), cette matrice représentera la matrice des taux de transition du processus agrégé lorque celui-ci sera markovien.
- Pour $l \in F \setminus \{0\}$, on notera $\sigma_l$ le système linéaire $x_l \widetilde{A}_l = \widehat{A}_l$. C'est un système à $(M+1)$ équations et $n(l)$ inconnues. Toute solution $x_l$ de $\sigma_l$ vérifie $x_l 1^* = 1$. Par construction, pour tout $l \in F \setminus \{0\}$, $\sigma_l$ a au moins une solution qui est $R_l.v^{B(l)}$.
- Nous définissons maintenant les ensembles suivants

$\mathcal{C}^1 \stackrel{\text{déf}}{=} \{\ \alpha \in \mathcal{A}^T\ /\ R_l \alpha^{B(l)}$ est solution de $\sigma_l \quad \forall l \in F \setminus \{0\}$ t.q $\alpha_{B(l)} \neq 0\}$

et pour $j \geq 2$

$\mathcal{C}^j \stackrel{\text{déf}}{=} \{\ \alpha \in \mathcal{A}^T\ /\ \forall \beta = h(\alpha, (B(i_0), t_0) \ldots, (B(i_k), t_k))$ avec $k \leq j,\ i_k \neq 0, \ \ \beta \in \mathcal{C}^1\}$.

Le résultat suivant donne une expression de l'ensemble $\mathcal{C}_{\mathcal{M}}^T$ en fonction des polyèdres convexes $\mathcal{C}^j$.

**Théorème 5.9**

$$\mathcal{C}_{\mathcal{M}}^T = \bigcap_{j \geq 1} \mathcal{C}^j$$

**Démonstration.** Elle est basée sur l'équivalence suivante du Théorème 5.6

$$\alpha \in \mathcal{C}_{\mathcal{M}}^T \Longleftrightarrow \begin{cases} \forall l \in F \setminus \{0\} \text{ et } \beta \in \mathcal{C}(\alpha, B(l)), \\ R_l.\beta \ \ \widetilde{A}_l = \widehat{A}_l. \end{cases}$$

$\square$

## 5.2 Lien avec le cas discret

La technique d'uniformisation [4] consiste à construire une chaîne de Markov $U = \{U_n, n \in \mathbb{N}\}$ sur l'espace d'états $E$, de matrice des probabilités de transition $P$, et un processus de Poisson $\{N_t, t \geq 0\}$ indépendant de $U$ et de taux $\lambda$ tels que $X_t = U_{N_t}$, de la façon suivante.

- On choisit $\lambda \in \mathbb{R}$ tel que: $\lambda \geq max(-A(i,i), i = 1, ..., N)$
- On pose $P(U_{n+1} = j/U_n = i) = P(i,j) = \left( \begin{array}{ll} 1 + \frac{A(i,i)}{\lambda} & si\ i = j, \\ \frac{A(i,j)}{\lambda} & sinon. \end{array} \right.$

  On obtient alors la relation suivante liant les matrices $A$ et $P$:

$$P = I + A/\lambda \tag{17}$$

Pour plus de clarté, on introduit les notations suivantes.

- On note $unif(X)$ la famille de chaînes $(., P)$.
- Pour tout $i \in F$, la matrice $\widetilde{I}_l$ de taille $(n(l), M+1)$ est définie par

$$\widetilde{I}_l(i,j) = \begin{cases} 1 & \text{si } j = l \\ 0 & \text{sinon.} \end{cases}$$

- La matrice identité $(M+1) \times (M+1)$ est notée $\widehat{I}$.

**Lemme 5.10** *Pour tout $\alpha \in \mathcal{A}$, le processus $(\alpha, A)$ et la chaîne $(\alpha, P)$ ont la même distribution de probabilité quasi-stationnaire.*

**Démonstration.** En appliquant le Théorème 3.2 de la Section 3.2 au processus $(\alpha, A)$, nous savons que $\forall j \in F \setminus \{0\}$

$$\lim_{t \to \infty} \frac{\alpha_T e^{Gt} e_j^{*}}{\alpha_T e^{Gt} 1^{*}} = v(j) > 0,$$

où $v$ est l'unique vecteur de probablité solution de $vG = \rho v$ avec $\rho$ rayon spectral de la sous-matrice matrice $G$ de $A$ obtenue par la décomposition de la Section 3.2. Nous déduisons de (17) que la chaîne absorbante $(\alpha, P)$ admet le même ensemble d'états transitoires que $(\alpha, A)$, que la sous-matrice $M \times M$ de $P$, $Q$, relative aux transitions entre ces états et satisfaisant la relation $Q = I + G/\lambda$, est irréductible et apériodique. La chaîne $(\alpha, P)$ admet donc une distribution quasi-stationnaire d'après le Théorème 3.1.

On montre facilement que $\rho + \lambda$ est le rayon spectral de la matrice $\lambda Q = G + \lambda I$ et que $v$ est l'unique vecteur de probabilité satisfaisant

$$v(\lambda Q) = (\rho + \lambda) v.$$

On en déduit que $(\rho+\lambda)/\lambda$ est le rayon spectral de la matrice $Q$ et que l'on a

$$vQ = (\frac{\rho+\lambda}{\lambda})v.$$

Ainsi $v$ représente également la distribution de probabilité quasi-stationnaire de la matrice irréductible $Q$ représentant les probabilités de transition entre les états transitoires de la famille de chaînes de Markov $unif(X)$. □

On peut maintenant énoncer le lemme suivant:

**Lemme 5.11**

$$\widehat{P} = \widehat{I} + \widehat{A}/\lambda$$

**Démonstration.** Soit $l \in F \setminus \{0\}$,

$$\begin{aligned}\widehat{A}_l &= R_l.v^{B(l)}\widetilde{A}_l \quad \text{par la relation (16)}\\ &= R_l.v^{B(l)}\lambda(\widetilde{P}_l - \widetilde{I}_l) \quad \text{par la relation (17).}\end{aligned}$$

Ce qui nous donne en développant avec le lemme précédent:

$$\widehat{A}_l = \lambda(\widehat{P}_l - e_l)$$

c'est à dire:

$$\widehat{A} = \lambda(\widehat{P} - \widehat{I}).$$

□

Finalement, on obtient le résultat principal de cette section permettant de traiter le cas des processus à temps continu comme celui des chaînes à temps discret dans le cadre de l'agrégation faible.

**Théorème 5.12** *La famille de processus $X = (.,A)$ est faiblement agrégeable selon la partition $\mathcal{B}$ ssi la famille de chaînes uniformisées $unif(X)$ est faiblement agrégeable. Dans ce cas*

$$\begin{cases} \mathcal{C}_{\mathcal{M}}(X) = \mathcal{A}_{\mathcal{M}}(unif(X)),\\ agr(unif(X)) = unif(agr(X)).\end{cases}$$

**Démonstration.** La preuve est identique à [7, Th. 3.2] une fois que l'on a établit le lemme précédent. Nous nous contentons d'en décrire les principales étapes. Pour une famille de chaînes discrètes $(.,P)$, on définit tout d'abord les éléments suivants.
Pour $l \in F \setminus \{0\}$, $\widetilde{P}_l$ est la matrice $n(l) \times (M+1)$ dont l'élément $(j,k)$ est défini par:

$$\widetilde{P}_l(j,k) = P(\,n(0) + \cdots + n(l-1) + j, B(k)\,), \quad 1 \le j \le n(l), k \in F.$$

On note $\widehat{P}_0, \ldots, \widehat{P}_l, \ldots, \widehat{P}_M$ les $(M+1)$ lignes successives de la matrice des probabilités de transition définie par la relation (9). Et on note $\sigma_l$ le système linéaire $x_l\widetilde{P}_l = \widehat{P}_l$.

On démontre alors que

$$\forall j \geq 1 : \mathcal{C}^j(X) = \mathcal{D}^j(unif(X)).$$

où les ensembles $\mathcal{D}^j$ sont définis pour la famille de chaînes discrètes $(., P)$ par

$\mathcal{D}^1 \stackrel{\text{déf}}{=} \{ \alpha \in \mathcal{A}^T \ / \ R_l.\alpha^{B(l)}$ est solution de $\sigma_l, \forall l \in F \setminus \{0\}$ t.q. $\alpha_{B(l)} 1^* \neq 0\}$

et pour $j \geq 2$

$\mathcal{D}^j \stackrel{\text{déf}}{=} \{ \alpha \in \mathcal{A}^T \ / \ \forall \beta = g(\alpha, \ (B(i_0), 0), \ldots, (B(i_k), n_k))$ avec $k \leq j, i_k \neq 0 \quad \beta \in \mathcal{C}^1\}$

la fonction $g$ étant l'analogue de la fonction $h$ en temps discret (i.e. les instants réels $t_k$ sont remplacés par des entiers $n_k$). Puis on obtient le même type de résultat que le Théorème 5.9 pour notre famille de chaînes discrètes

$$\mathcal{A}^T_{\mathcal{M}} = \bigcap_{j \geq 1} \mathcal{D}^j.$$

Enfin, avec le Théorème 5.9, on conclut que $\mathcal{C}^T_{\mathcal{M}}(X) = \mathcal{A}^T_{\mathcal{M}}(unif(X))$. La seconde égalité du théorème traduit simplement le Lemme 5.11.

□

# Bibliographie


[1] J.N. Darroch and E. Seneta. On quasi-stationary distributions in absorbing discrete-time finite Markov chains. *J. Appl. Prob.*, 2:88–100, 1965.

[2] J.N. Darroch and E. Seneta. On quasi-stationary distributions in absorbing continuous-time finite Markov chains. *J. Appl. Prob.*, 4:192–196, 1967.

[3] J.G. Kemeny and J.L Snell. *Finite Markov chains.* Springer-Verlag, New York Heidelberg Berlin, 1976.

[4] S.M. Ross. *Stochastic Processes.* John Wiley and Sons, 1983.

[5] G. Rubino and B. Sericola. *Agrégation d'états dans les processus markoviens.* Rapport INRIA RR-858, Avril 1988, Campus de Beaulieu, 35042 Rennes Cedex, France.

[6] G. Rubino and B. Sericola. A finite caracterization of weak lumpable Markov processes. Part I: The discrete time case. *Stoch. Proc. and Appl.*, 38:195–204, 1991.

[7] G. Rubino and B. Sericola. A finite caracterization of weak lumpable Markov processes Part II: The continuous time case. *A paraître dans Stoch. Proc. and Appl.*

[8] G. Rubino and B. Sericola. On weak lumpability in Markov chains. *J. Appl. Prob.*, 26:446–457, 1989.

[9] E. Seneta. *Non-negative matrices and Markov chains.* Springer-Verlag, 1981.

## LISTE DES DERNIERES PUBLICATIONS INTERNES PARUES A L'IRISA

PI 661 REACHABILITY ANALYSIS ON DISTRIBUTED EXECUTIONS
Claire DIEHL, Claude JARD, Jean-Xavier RAMPON
Juin 1992, 18 pages.

PI 662 RECONSTRUCTION 3D DE PRIMITIVES GEOMETRIQUES PAR VISION ACTIVE
Samia BOUKIR, François CHAUMETTE
Juin 1992, 40 pages.

PI 663 FILTRES SEMANTIQUES EN CALCUL PROPOSITIONNEL
Raymond ROLLAND
Juin 1992, 22 pages.

PI 664 REGION-BASED TRACKING IN AN IMAGE SEQUENCE
François MEYER, Patrick BOUTHEMY
Juin 1992, 50 pages.

PI 665 CORRECTNESS OF AUTOMATED DISTRIBUTION OF SEQUENTIAL PROGRAMS
Cyrille BARREAU, Benoît CAILLAUD, Claude JARD, René THORAVAL
Juin 1992, 32 pages.

PI 666 AGREGATION FAIBLE DES PROCESSUS DE MARKOV ABSORBANTS
James LEDOUX, Gerardo RUBINO, Bruno SERICOLA
Juillet 1992, 30 pages.

PI 667 MODELES D'EVALUATION DE LA FIABILITE DU LOGICIEL ET TECHNIQUES DE VALIDATION DE SYSTEMES DE PREDICTION : ETUDE BIBLIOGRAPHIQUE
James LEDOUX
Juillet 1992, 76 pages.

PI 668 TWO COMPLEMENTARY NOTES ON SKEWED-ASSOCIATIVE CACHES
André SEZNEC
Juillet 1992, 10 pages.

PI 669 PARALLELISATION D'UN ALGORITHME DE DETECTION DE MOUVEMENT SUR UNE ARCHITECTURE MIMD
Fabrice HEITZ, Sergui JUFRESA, Etienne MEMIN, Thierry PRIOL
Juillet 1992, 34 pages.

ISSN 0249-6399